

## CAFE CBA Team Response to:

# UNICE CONCERNS WITH KEY ASPECTS OF CAFE CBA METHODOLOGY

05 May 2005

**Fintan Hurley**, IOM, Edinburgh  
**Mike Holland**, EMRC  
**Paul Watkiss**, AEA Technology  
**Alistair Hunt**, University of Bath

### Introduction and Summary of our response

UNICE sent the communication identified above to the European Commission on April 13<sup>th</sup> 2005. This note provides the response of the CAFE CBA team to that document. For clarity, we include the full text of the UNICE letter *in italics*, to distinguish it from our response, and make some other minor format changes to the headings, for clarity.

We have read and considered the comments made by UNICE very carefully. We appreciate the care taken by UNICE to present their views in detail, and we respond in them in detail below.

Very briefly, our response is that

- i. UNICE suggests that issues of uncertainty be described and highlighted more fully within the CAFE CBA methodology. This is being done within Volume 3 of the CBA series of reports, which has been submitted to the Commission recently for final approval.
- ii. Other views expressed by UNICE relate almost entirely to potential for overestimation of benefits in the methods defined for the CAFE benefits analysis.
- iii. The views expressed by UNICE go against the current consensus in several areas, and in particular:
  - a. The assessment of the evidence by various World Health Organisation (WHO) review and working groups,
  - b. The assessment of the evidence by the independent peer review team for CAFE CBA.
- iv. The views expressed by UNICE also under-estimate the strength of the evidence on which quantification of morbidity is based.
- v. After due consideration, we consider that the methodology as peer reviewed and then revised is an appropriate and robust methodology.
- vi. Our detailed considerations follow.

### UNICE Concern:

*This letter addresses some key concerns relating to the CAFE Cost Benefit Assessment (CBA) health methodology now that the Final Methodology Volume 2 Health Impact Assessment has been made available.*

### Response:

We appreciate the care and detailed work that underlies UNICE's response to the current methodology.

Several important issues are raised. We would however like readers of UNICE's letter to remember that UNICE correctly describes this as a letter stating UNICE's *concerns*. Two things follow:

- It is not a review of the CAFE CBA methodology. A review would identify perceived strengths as well as perceived weaknesses. Generally the letter doesn't focus on strengths, though various industry representatives have been complimentary on aspects of the methodology on earlier occasions, both in verbal comments at stakeholders' meetings, and in written comments.
- UNICE's concerns refer almost entirely to where health effects may be over-estimated. Except for the effects of coarse particles they do not refer to possible under-estimation, e.g. by the CAFE CBA methodology not being able to quantify an effect of long-term exposure to PM on development or acceleration of chronic cardiovascular disease, though an effect is very highly likely, but there were no suitable studies for quantifying it.

**Concern:**

*Our general comment is that there has been a breakdown in what we consider to be an important aspiration for CAFE - stakeholder engagement and transparency. We are particularly concerned that there was no opportunity for stakeholders to be involved or comment in the process that resulted in the selective inclusion of a number of the peer reviewer's comments into the core CAFE CBA methodology.*

**Response:**

The draft methodology was presented and discussed in July 2004, with participation of all major stakeholders of the CAFE programme, including UNICE. The stakeholders views were provided in the workshop and the supporting written comments were carefully considered by the CBA team in the revision of the CBA methodology.

The peer review of the draft methodology that took place in September and October 2004 had the objective, according to the Technical Annex of the assignment, "to have reviewed the methodology of cost-benefit analysis and to what extent it is fit for purpose for policy application in the Clean Air for Europe Programme. This review is independent from the Commission or stakeholders and is using a scientific approach." The last point is specifically important since the objective is that the methodology should be based on sound science and not tendentious or biased, and systematically trying to 'talk up' or 'talk down' the effects of air pollution on health.

The peer reviewers were selected on the basis of their recognised expertise in the field. For example Bart Ostro is possibly the best-known expert internationally on quantifying the health effects of outdoor air pollution.

In general on health issues we have adopted the reviewers' comments. The main exceptions are where there was an established recent WHO position, from its recent review work for CAFE, from its meta-analysis of selected endpoints from studies in Europe, and from the Task Force on Health (TFH) of the WHO-UNECE Convention on Long-Range Transboundary Air Pollution (CLRTAP). We decided to maintain consistency with the WHO position (unless there was important new evidence not considered by WHO) (i) because the WHO position was reached as the result of discussion and evaluation by one or more groups of established air pollution experts and (ii) to maintain consistency of the health evaluations within CAFE as a whole.

**Concern:**

*Our main concerns are:*

- *all aspects of morbidity assessment;*
- *use of the VSL (Value of Statistical Life) metric rather than the VOLY (Value of One Life Year) metric;*
- *use of mean valuation values rather than median values;*
- *use of a single exposure response function and ignoring cutpoints; and*
- *the misleading manner that health endpoints were conveyed in a recent non-paper.*

*In our opinion all of the above are likely to have very significant implications for policy setting, not only for CAFE and the Thematic Strategy on Air Pollution, but also IPPC regulation through the reference to the CAFE CBA methodology included in the IPPC Economics and Cross Media BREF.*

*Please find below more detailed information on these concerns and constructive suggestions on improving the communication of the CBA results.*

**Response:**

For ease of reference we will number the five issues following, raised by UNICE. A short response to various summary points of UNICE's is followed by more detailed response to comments on selected specific studies (generally, those that gave highest impacts).

## **1. Assessment of morbidity**

### **Concern:**

*We are concerned with the accuracy of the morbidity benefits currently appearing in the CAFE CBA. Early in the process of developing the CBA methodology, assessment of morbidity endpoints was viewed by the CBA project contractor as very challenging*

### **Response:**

It is correct to say that this work is challenging, and so the development of the methodology for morbidity assessment has required substantial work by the CBA team. However (i) it builds on previous published work by ExternE and other HIA teams, which UNICE and others have commented on in the past; and it updates and extends that work by (ii) using more up-to-date studies for the concentration-response functions (CRFs), including by WHO and by APHEIS, and (iii) by more comprehensive assessment of background rates of morbidity. The sources and assumptions are based on published findings and well established methodology and are well documented in the methodology papers of the CBA, especially Volume 2.

### **Concern:**

*and not possible to any reasonable degree of accuracy*

### **Response:**

We did not say that. We always expected to achieve a 'reasonable degree of accuracy', if by this is meant 'good enough to be used for policy purposes'. This expectation was based on our experience of ExternE and of other HIA studies, and on the fact that, we always. The CAFE CBA team has quantified morbidity impacts as part of the European Commission's ExternE programme for at least 10 years now. In addition, we knew that further extensive review work and meta-analyses had been carried out by WHO and others, work which we knew would help improve the quantification – see previous answer.

Of course we recognise that the assessment of morbidity is linked with some uncertainties that have to be addressed as part on the sensitivity and uncertainty analysis. But the methodology and the assumptions are not so large as to make quantification impossible. Our guideline has been "Is the overall quantification more accurate by including this endpoint than by omitting it?". On that basis we expected to, and did, include a range of morbidity endpoints, as indeed do other major quantification projects.

### **Concern:**

*The contractor cited a variety of scientific problems including:*

- *a very limited number of reliable exposure response functions, especially European specific functions* **Response:** this varies by endpoint, but is true for some endpoints, including some which have proved to be the most influential;
- *lack of reliable baseline rates for morbidity effects needed to accurately determine the fraction attributed to air pollution exposure; and*
- *limited data on valuation of morbidity effects.*

*However, it appears that largely due to the comments of a single reviewer, morbidity endpoints now receive high visibility in the current CBA.*

### **Response:**

The CAFE CBA team thought from the beginning that morbidity endpoints would have 'high visibility' within the CBA. Indeed, we consider that they are highly visible within the July 2004 draft. What we did not know, until we did the work, is whether or not they would have 'high importance'. It turns out that they have quite high importance.

### **Concern:**

*This is despite all of the original reasons for not including morbidity effects remaining valid.*

### **Response:**

See earlier – it is inaccurate to describe the uncertainties listed earlier as 'reasons for not including morbidity effects', and certainly inaccurate to ascribe that view to the CBA team.

**Concern:**

No new information was introduced during the peer review on this matter.

**Response:**

Generally, this is true. However:

- i. whether on balance the CBA is stronger or weaker by including particular endpoints is a matter of judgement, and the comments of the review team were relevant though not determining; and
- ii. the CBA team extended its evaluation of morbidity considerably in revising the July 2004 draft following reviewers' comments. In particular, we strengthened the evaluation of background rates. In the July draft we commented on likely sources of information. By November 2004 these had been assessed and included.

Note: Already in February 2004 (Methodology, Issue Two, Section 1.3.4) we stated (emphasis added now) that: "Our aim is that morbidity effects are quantified, but that the uncertainties are fully explained and where practicable quantified or explored through sensitivity analysis. Aggregation of benefits can and will include specific impact pathways, or not, according to their reliability – see later, under uncertainty analysis. **The alternative is to omit from the quantification some of the impact pathways that are widely believed to reflect some of the benefits of reducing air pollution, purely on the grounds that sufficiently reliable baseline data are unavailable.** This may in some instances be necessary. **However, we see it as the option of last resort.** Within the context of a CBA it is important to know at least roughly the morbidity benefits of reducing air pollution, and we aim to provide that information – partly to assist in evaluating scenarios and options; partly to help identify what are the information gaps that have most influence on the final results of the HIA, and to identify how these might best be filled." Also, in the consultation draft of July 2004 it was clearly stated what was to come in the final methodology, and also indicated that Restricted Activity Days from exposure to PM and ozone would likely be the most important effect.

To summarise: From the outset, i.e. the first stakeholders' meeting of December 2003, the thrust of our work has been to quantify as much as possible, while recognising that necessarily there are uncertainties. The quoted Section is reproduced, with minor editing, again as Section 1.3.4, of the present Volume 2. We do not understand how UNICE got the idea that mostly we would not quantify morbidity.

**Concern:**

We also note that many of the key exposure response functions used to assess morbidity effects come from studies published by the same peer reviewers who strongly advocated their use.

**Response:**

This seems to be presented as a criticism. It is misleading to say that Ostro published the studies that provide 'many of the key functions', though he did publish some. Presumably he knows the strengths and limitations of his own work! We reject the implication that there is something wrong in using studies of which Bart Ostro is a main or co-author, as if his recommendation of these studies was to publicise his own work.

These problems are discussed further in the attachment to this letter, along with suggestions for improving the accuracy of the morbidity calculations.

**2. Use of the VSL (Value of Statistical Life) rather than the VOLY (Value of One Life Year)**

**Concern:**

*The Newext researchers, as well as by Rabl et al., have pointed out that VSL is not the correct metric for use in the air pollution context. This is because the relevant health effect is a change in life expectancy. This is further supported by the fact that cohort studies analyzing air pollution effects can only reveal a life expectancy change, not a change in the number of deaths.*

**Response:**

This is now widely believed – and I have in the past said so myself – but it is strictly not true.

- The coefficients from the cohort studies represent a relative risk or % change in *mortality hazard rates*, which in effect are age-, sex- and race-specific death rates.

- To begin with, these coefficients were used to estimate changes in the numbers of deaths in a 'target' population for quantifying impacts
- Then, life table methods were developed, following an initial paper by Brunekeef (1997). These give results which are most naturally expressed in terms of changes in life expectancy. The use of life tables was developed most notably within ExternE by what became the CAFE CBA team. Since 1997-98, the practice of the CAFE CBA team has been to quantify using life tables, giving results in terms of life expectancy.
- However, it is possible to give estimates of numbers of deaths also. In practice, it gives the same results as would be given from the 1<sup>st</sup> year's results of using a life table. These 1<sup>st</sup> year results are then used for each succeeding year. This is an approximation (strictly speaking, incorrect) but there is not an established view on its reliability.
- We recognize that, largely for reasons of linkage with economic valuation, there is a strong case that mortality to be evaluated in terms of numbers of deaths as well as changes in life expectancy. Many economists, including some CAFE stakeholders, disagree with the VOLY approach, pointing out that this is not how people value their own risks of dying. The reviewers supported this position. Although a number of people are coming round to accept the VOLY approach it would be wrong to say that there is a decisive majority in support of it. Reflecting this, it is correct from a pragmatic perspective to use both VSL and VOLY based methods.
- The view of the CAFE CBA team was and is that all quantification involves approximations, and that the approximations involved in estimating deaths from long-term exposure are not so serious that this should not be done, and so we have included it.
- In fact, the choice of approach does not matter much. The impression often given is that VSL based results are many times higher than VOLY based results. Volume 3 and the CBA team's reports on scenario analysis demonstrate that this is not the case.
- Volume 2 gives all this in greater detail.
- To summarize, it is not true that results from cohort studies cannot be expressed in terms of numbers of deaths, at least approximately.

It may help to summarize briefly the debate among economists regarding the use of VSL and VOLY.

### The case "for" and "against" the use of VSL

#### **Arguments for VSL:**

1. **Convenience.** The VSL is a convenient concept which involves dividing some estimate of the willingness to pay for a mortality risk reduction by that risk reduction. This measure is convenient because it can be multiplied by the statistical deaths averted by a policy to arrive at the benefits of that policy for this health endpoint.

2. **Existing experience in measurement and use.** The VSL measure is the traditional one. It is based on a huge literature of original studies and meta-analyses. It has been repeatedly endorsed by peer review panels in the U.S.

3. **Complementarity with physical measurement for effects of short-term exposures (from time series studies):** VSL can be applied from estimates of lives lost taken directly from the epidemiological literature, which typically offers estimates of deaths averted by or attributable to a change in pollution. However, this would only cover the damage due to short term exposure and not the other effects due to long term exposure and chronic effects.

#### **Arguments against VSL:**

1. **(Lack of) Appropriateness.** Much of the supporting literature is for a context not related to environmental policy -- wage premiums paid for risky jobs, involving people of better than those at risk from pollution, in the prime age groups, and taking voluntary risks of mostly immediate death.

2. **Limited complementarity with physical measurement for effects of long-term exposures (from cohort studies):** This has been discussed, above.

## **The case “for” and “against” use of VOLY**

### **Arguments for VOLY**

1. **Convenience.** The VOLY measure is convenient because it can be multiplied by the statistical life-years averted by a policy to arrive at the benefits of that policy for this health endpoint.
2. **Appropriateness.** Most of the deaths delayed from environmental policy are to the elderly and people with severe cardiovascular diseases and to treat deaths in these two groups as equivalent for valuation purposes seems inappropriate because so many fewer life-years are lost when the elderly die.
3. **Complementarity with physical measurement for effects of long-term exposures (from cohort studies):** This has been discussed, above.

### **Arguments against use of VOLY**

1. **(Lack of) Existing experience in measurement.** The VOLY measure does not have the lineage enjoyed by the VSL since the VOLY literature is less extensive, involving only a few studies that directly ask for the Willingness To Pay (WTP) for additional life expectancy. Unlike the VSLs, which are computed from estimates of the WTP for risk reductions using hedonic wage models, stated preference surveys, and consumer behaviour surveys, the VOLYs have been computed mainly through computational adjustments of existing VSLs. Hence there is a need to further develop the methodology and obtain independent new information of WTP, specific for cause and age etc., expressed in VOLYs.
2. **Limited complementarity with physical measurement for effects of short-term exposures (from time series studies):** Use of VOLYs requires 'conversion' from deaths to associated life-years and, while judgements on this can be made for the short term effects, there is no direct evidence – see CAFE CBA, Vol 2.
3. **Inconsistency between VOLY and VSL:** One interpretation of the use of VOLYs is that they should sum to produce a VSL. However, in reality it is evident that expected remaining lifetime is not the sole driver of a VSL, as this interpretation would suggest. This is evident because VSL expressed at different ages does not fall from birth, as it would if only expected remaining lifetime drove this estimate. Instead it rises from birth, generally peaking in the 40's during one's high earning period, and falling thereafter, showing that other factors such as income level and number of dependents are also important.
4. **Egalitarian arguments against use.** U.S policy guidance effectively prohibits use of a VSL lower for the elderly than for younger people.<sup>1</sup> While direct translation of this prohibition to the VOLY would imply simply that VOLYs could not vary by age group, the spirit of the debate is clearly that VOLYs should not be used in policymaking because this would discriminate against the elderly, who have fewer life-years to live.

#### **Concern:**

*This is why the CAFE CBA contractor initially proposed to use only VOLY as a valuation metric, and industry fully supported this approach. We do not agree with the comments of one of the peer reviewers that VSL should also be considered.*

#### **Response:**

This is covered, above. Moreover, both estimates are provided in the analysis, giving a bracket of valuations.

#### **Concern:**

*As a further remark, noting that VSL is an irrelevant metric within the air pollution context*

#### **Response:**

No: see above

---

<sup>1</sup> [http://thomas.loc.gov/cgi-bin/bdquery/z?d108:HZ00338:](http://thomas.loc.gov/cgi-bin/bdquery/z?d108:HZ00338)

**Concern:**

*and noting that mean values for WTP are meaningless*

**Response:**

We disagree – see later

**Concern:**

*we strongly disagree with the current practice of the CAFE CBA team to quote four results for the benefits estimates (VOLY and VSL both with mean and median) and calling this a sensitivity analysis - three out of these four results are in reality irrelevant and therefore this does not represent any kind of meaningful sensitivity analysis.*

**Response:**

We have described, above, why we think that both VSL and VOLYs should be used. We describe below why we think that both median and mean are relevant. Hence we do not agree that only the median value of VOLY should be used for the analysis of policy options – the four values given are a legitimate sensitivity analysis. Other aspects of sensitivity, e.g. use of a range of values for the relative risk (where the core is 6% increase in mortality hazard rates, per 10µg/m<sup>3</sup> PM<sub>2.5</sub>) form the basis of Volume 3 of the CAFE CBA series, [which has been submitted to the Commission recently for final approval](#). Some specific points are discussed further, below, in our comments on UNICE's Attachment 2.

**3. Use of arithmetic mean rather than median values**

**Concern:**

*Several authors, including Rabl et al. have pointed out that, due to the underlying statistical distribution of the WTP values found in Contingent Valuation studies, the arithmetic mean is not a robust statistic to characterize the WTP properties. It is far too sensitive to model assumptions concerning the WTP probability distribution, which usually has a few very large numbers which have a disproportionate impact on the arithmetic mean. Authors agree that the median is a much more robust statistic, although of course an analysis taking the full distribution into account is the best approach. Again, the CAFE CBA contractor initially realized this as well. It was again based on the comments of one of the peer reviewers that the mean value was reintroduced.*

**Response:**

We agree with the peer review comments that “there is no consensus on the appropriate statistic for describing the average sample response”.

- It seems that the distribution of WTP values is similar to the distribution of income, i.e. very skewed. Under those circumstances, the median is more robust than the mean, i.e. less sensitive to occasional extreme values. Also, it is a conservative measure, being lower than the mean because, in part, the WTP responses are bounded at zero but not bounded from above.
- However, robustness isn't everything – we also need relevance. An argument for use of the mean is that it is conceptually the appropriate measure for use in CBA because it fully summarizes the heterogeneity of values in the sample. If the decision-maker wishes to make a decision based on efficiency criteria, (as the CAFE CBA requires), then the mean is, strictly speaking, the more appropriate measure.
- It seems then that the choice is between a more robust estimator of a less relevant measure, or a less robust estimator of a more relevant measure. Under those circumstances it seems reasonable to present both.
- The uncertainty assessment investigates the distribution on valuation metrics in detail. However, it is still useful to present numbers so that people can see what happens when you use mean or median numbers.

**4. Key Exposure Response Functions and Cut-points**

**Concern:**

*As described in detail in Attachment 2, we recommend use of alternate scientifically-based exposure response functions for chronic PM<sub>2.5</sub> mortality and acute ozone mortality, with appropriate cutpoints as employed by USEPA. For chronic PM mortality, the data support consideration of point values of 4%*

and 1% per 10 ug/m<sup>3</sup> for all cause mortality with a cutpoint in the range of 7-12 ug/m<sup>3</sup>. For acute ozone mortality, the data support a point value of 0.2% /ug/m<sup>3</sup> with a cutpoint of 50-60 ppb.

**Response:**

CAFE CBA used recommendations from WHO (TFH, WHO-UNECE). We elaborate on these choices in our response to UNICE Attachment 2, though without re-iterating all the arguments which underlie the positions that WHO adopted on these and associated issues.

**Concern:**

*Also, we suggest that the CAFE programme explore the use of distribution functions for both exposure response and valuation, which capture the full range of risks and uncertainties associated with health impact assessments. This point was highlighted in the recent SCHER opinion on 'New evidence of air pollution effects on human health and the environment', which stated that the 'results from health risk assessments should deal with uncertainties and should at least be presented with relevant uncertainty margins.'*

**Response:**

We agree; this issue is being addressed. Methods and some results will be presented in CAFE Methodology, Volume 3.

## **5. Communication of CAFE CBA results**

**Concern:**

*We are very concerned that the morbidity endpoints for which quantification was viewed as speculative in the CBA methodology before the peer review (and therefore proposed for inclusion in the sensitivity analysis) are now included in the core analysis and communicated along with the other endpoints without any distinction or recognition of the increased uncertainties.*

**Response:**

There are two issues here:

- i. Whether to include as core, sensitivity, or not at all. This has been addressed under 1, above
- ii. How to communicate the results: we agree with the general point, that the reporting of results should include some better signalling of the basic assumptions, uncertainties and relative accuracy, including in how the tables are presented.

**Concern:**

*These health endpoints include various forms of restricted activity days, respiratory medication use in children and adults, and lower respiratory symptoms and cough in children and adults. Infant mortality, previously described as too speculative for accurate assessment, is also now included in the core analysis.*

**Response:**

We do not recall describing infant mortality as "too speculative for accurate assessment". We expected to include it, as a core quantification, and signalled this from an early stage of the CBA process. We strongly recommend that the CAFE CBA programme clearly separate out these endpoints in reporting, rather than combine them. We agree (see below) that we should signal better the uncertainties associated with various endpoints. This will be addressed in Volume 3 which, as noted above, has been submitted to the Commission recently for final approval. See also some of the comments, following.

**Concern:**

*Finally, we are also concerned with how the morbidity benefits and the mortality benefits are being presented and communicated under the CAFE programme. Essentially, results based on retrospective ecological-based studies with high uncertainty are being projected into the future and communicated as actual benefits that can be tracked and stewarded to fine detail. This may be the case for monthly safety statistics, but it will not be possible to track any actual changes in incidence of mortality or morbidity resulting from future reductions in air pollution.*

**Response:**

We agree that that the estimated benefits should be presented as estimates, not as exact values that are certain to occur. In the more polluted areas a drastic change of air pollution levels could result in clearly measurable changes of improved health, particularly if the study design is set up for that purpose. Such changes of health effects due to sustained changes in air pollution have been noted in the past, notably the 1990 ban on domestic coal burning in Dublin (Clancy et al., 2002; Lancet; 360: 1210-14) and the 1990 ban on sulphur in fuel in Hong Kong (Hedley et al., 2002; Lancet; 360: 1646-52). To a lesser extent, temporary pollution reductions (notably the temporary closure 1986-87 of the Geneva steel mill in Utah Valley – Pope et al., 1992; Arch Environ Health; 47: 211-217) have had marked effects on mortality and morbidity also. On other cases, we will not be able to link actual changes in pollution reliably with actual changes in mortality and morbidity. This is because, in general, there are so many factors influencing mortality and morbidity that changes in one risk factor may not show clear effects against the background of variation caused by other factors.

**Concern:**

*The Commission, Parliament, Council and public should not be given the impression that this is the case. Rather, at this critical stage of the CAFE Thematic Strategy deliberations, they should be informed that it will not be possible to track and steward actual disease trends using the existing studies and that there is a high degree of uncertainty in the future benefit projections. This uncertainty needs to be clearly indicated so that informed decisions can be reached. To help illustrate how the tables could be improved, we have provided, see Attachment 3, an edited version of the table provided in a recent CAFE non-paper along with some appropriate notations that capture some of the key uncertainties.*

**Response:**

We agree with the principle of signalling different degrees of certainty within a results table, and have done so within ExternE. (Of course there is an unavoidably subjective element to the assessment of uncertainty.) The Table presented in Attachment 3 has some helpful presentation ideas in terms of methods; for example, rounding of numbers, use of footnotes. The exclusion of results from many of the quantified endpoints does not however reflect the CAFE CBA team's assessment of the relevant uncertainties.

In conclusion, the CAFE CBA team recognizes the careful work by UNICE underlying this statement of concerns and we hope that this response may go some way towards putting those concerns to rest.

## Attachment 1

### DETAILED COMMENTS ON EVALUATION OF MORBIDITY EFFECTS

#### BRONCHITIS

**Concern:**

*We have multiple concerns with how this endpoint is being evaluated. **First, the entire impact of the contribution of particulate matter on the incidence of bronchitis is inappropriately attributed to the fine particle fraction.** This defies well-known biological facts concerning the etiology of bronchitis and highlights the need for clinical input into the CAFE CBA. It is very well known that bronchitis is primarily a disease of the upper respiratory tract. Therefore, coarse particles, which deposit in the upper respiratory tract, are much more likely to contribute to the etiology of this disease. Fine particles, which deposit in the lower respiratory tract, are not expected to contribute to the incidence of bronchitis. Therefore, it is not biologically appropriate to convert the morbidity function from a study using PM<sub>10</sub> to PM<sub>2.5</sub>. Rather, as described in our previous comments, for bronchitis, a separate benefits analysis for PM<sub>10</sub> or other coarse particle metric such as PM<sub>2.5-10</sub> or TSP should be provided. It is critical to note that in the study by Abbey et al, a stronger relationship was observed for TSP, the actual metric of particle exposure used, than for PM<sub>10</sub> or PM<sub>2.5</sub>. In our view, valuation of bronchitis attributed to PM<sub>2.5</sub> should not be performed under the CAFE programme.*

**Response:**

We agree that the coarse fraction is relevant; we do not agree that the fine fraction is irrelevant.

Volume 2 gives a C-R function and associated impact function in PM<sub>10</sub> (Sec 8.3.3 and 8.3.4) and one in PM<sub>2.5</sub> (Sec 8.3.5 and 8.3.6). The preferred impact function, i.e. the function to be used in the core analysis, is the function in PM<sub>10</sub> (Sec 8.3.7).

The problem arises because RAINS does not give modelled pollution data in PM<sub>10</sub>.

**Concern:**

***This stated, we are reasonably certain that the authors of the CBA have not used an appropriate method to convert the exposure response function (ERF) based on PM<sub>10</sub> to an ERF based on PM<sub>2.5</sub>.** The authors of the CBA take the attack rate for chronic bronchitis based on PM<sub>10</sub> and adjust it to get the results for PM<sub>2.5</sub>. In our view, they should have adjusted the attack rate to get the PM<sub>2.5</sub> fraction, then taken out PM<sub>2.5</sub>. Assume that 54% of the PM<sub>10</sub> attack rate is due to PM<sub>2.5</sub> (a conversion factor of 1.54 is used). As such, 54% of the attack rate based on PM<sub>10</sub> (which is 7%) becomes a 3.8% attack rate. So, if 54% of the PM<sub>10</sub> exposure is PM<sub>2.5</sub>, this means a 10 ug/m<sup>3</sup> PM<sub>10</sub> ERF equates to 5.4 ug/m<sup>3</sup> PM<sub>2.5</sub>. Final adjustment should be 5.4 ug/m<sup>3</sup>/3.8% attack rate = 0.7% adjusted attack rate for PM<sub>2.5</sub>. This compares with the authors adjusted rate of 1.07%.*

**Response:**

There are two ways of looking at this conversion:

- i. The ideal is to quantify in PM<sub>10</sub>. Our goal, therefore, is to assess the impacts that would be estimated if we had pollution data in PM<sub>10</sub> and could work in this metric directly. Without the PM<sub>10</sub> data we need to carry out the calculations in PM<sub>2.5</sub>. The issue then, is: what CRF in PM<sub>2.5</sub> will, when linked with pollution data in PM<sub>2.5</sub>, give us the best estimates of the impacts we would have estimated if we had been able to work with PM<sub>10</sub>? That is what we have tried to estimate.
- ii. The second way of looking at it is to estimate what effects are associated with the PM<sub>2.5</sub> fraction alone, assuming that the effects of PM<sub>2.5</sub> (compared to PM<sub>10</sub>) are directly proportional to the mass of PM<sub>2.5</sub> as a % of the mass of PM<sub>10</sub>. This is what UNICE proposes.

In our view, the existing CAFE CBA approach is the more appropriate. Indeed, it is more consistent with UNICE's view of underlying mechanisms, which should take account of the coarse fraction.

The key to this is that we wish to quantify in PM<sub>10</sub>, but we have to do this in an indirect way – via a quantification in PM<sub>2.5</sub> – as a computational device, because we do not have modelled PM<sub>10</sub> data. We are not attributing all the effect to the PM<sub>2.5</sub> fraction only, in some biological sense.

**Concern:**

**We are concerned that the ERFs used were not even for the air pollutant under consideration.** Since monitoring of both  $PM_{10}$  and  $PM_{2.5}$  was very limited in California before 1986, Abbey et al. used data for Total Suspended Particulate (TSP) to derive estimates for  $PM_{10}$  and airport visibility records to derive estimates for  $PM_{2.5}$ . This very approximate approach to exposure estimates seriously jeopardized the findings from the study.

**Response:**

We agree that the results in  $PM_{10}$  and  $PM_{2.5}$  are derived, not directly measured. This is not unusual in studies using historical data. This contributes to uncertainty but it does not 'seriously jeopardize' the results of the study. In fact, errors in conversion will lead to mis-classification of exposures, and mis-classification of exposures tends to result in *under-estimation*, not over-estimation, of the associated CRF.

**Concern:**

**We are concerned that the assessment of bronchitis is based on a single study (ASHMOG Abbey et al., 1995a) for which the result was not statistically significant at the 5% level.** We do not believe that causality can be established based on the results of a single ecological epidemiology study.

Further, we question the accuracy of an ERF based on a single study result.

**Response:**

In order to best estimate annual benefits, we have examined PM and new cases of chronic bronchitis and we agree that possible CRFs are available only from the AHSMOG cohort. However,

- a. We establish/ infer causality based on the evidence as a whole, not just for the health endpoint under consideration. This principle is described, e.g. in the discussion of 'coherence' in the report of the AIRNET Working Group on Risk and Impact Assessment (Working Group 4) (available at <http://airnet.iras.uu.nl/>). That coherence includes that long-term exposure to PM is associated with bronchitis in other studies, e.g. Schwartz (1993; Environ Res; 62: 7-13) and also with reduced lung function. It is then to be expected that new cases of bronchitis are also associated with long-term exposure to PM.
- b. The AHSMOG study has produced many papers giving CRFs, with different ways of assessing health effects and exposures. These various papers/ analyses give similar results. This is not as strong as corroboration from different studies. Neither is it so weak as results from a single paper only.

**Concern:**

*In particular, we are concerned with the accuracy of the adjustment for smoking in this study, which is the major contributor to the overall incidence of bronchitis. The authors make the assumption that no smoking occurred in the cohort of seventh day Adventists. The higher lung cancer rates for males versus females in this study raises concern for this assumption. Further, 15% of the subjects in the study had smoked prior to 1977 and were then assumed to stop smoking when they became seventh day Adventists. Thirty percent of the study subjects lived with a smoker, and 42% had worked with a smoker.*

**Response:**

The Seventh Day Adventists were chosen for study because of a lifestyle/ belief system that does not include smoking. There is no doubt that smoking is less of a factor for this study group than for the population generally.

What if, as suggested, there is nevertheless an effect of smoking, and, perhaps, some under-reporting? Then, although smoking is clearly a risk factor for development of bronchitis, it does not follow that it is a confounder of the relationship between PM and bronchitis.

- In coalminers, for example, there is an effect of smoking on bronchitis and an effect of exposure to (coalmine) dust. However, these two effects are independent – the estimated effect of dust is basically the same, whether smoking is included in the model, or not.
- Smoking would be a confounder if, and only if, smoking habit was also related to exposure to PM; and I see no good reason why the (unreported) heavier smoking should occur in those who were more heavily exposed to outdoor PM, once age has been included in the analysis (as indeed it was).

- Smoking is a risk factor for mortality in the ACS study but adjustment for smoking did not alter the estimated effects of PM on mortality in the ACS cohort.

Most likely this is not a serious flaw.

**Concern:**

*Further, the sheer size of the risk due to PM air pollution, which is essentially the same size as the background rate attributed to all other factors, raises more suspicion.*

**Response:**

It is less easy to judge what is 'suspicious' in a population consisting mostly of non-smokers.

**Concern:**

*Finally, we find it questionable to use an ERF based on a result that was not statistically significant at the 5% level.*

**Response:**

Several closely related CRFs from the AHSMOG study are statistically significant, or close to it, at the 5% level. We have looked on statistical significance at the 5% level as a guideline, not as a strong decision rule. This is consistent with standard good practice in evaluating epidemiology studies.

**Concern:**

*This brings into the question the concept whether the findings from a single study are robust enough to conclude in the primary or core portion of a CBA. In our view, the approach to use a single study was influenced by the reviewer of the CBA methodology, who advocated use of a study for which he is a co-author. It is unlikely that such an approach would normally be justified and we do not understand why the authors of the CAFE CBA accepted this recommendation of the reviewer seemingly without question.*

**Response:**

Our inclusion of bronchitis in the core analyses was not unduly influenced by the comments of the review team. We have included bronchitis in core analyses previously, in Externe studies.

**Concern:**

***We are concerned that using data from California during the period of 1966-1988 when air pollution was high, will likely resulting in an inflated ERF.*** *It is interesting to note that the authors of the CBA choose to partially justify inclusion of this endpoint based on reference to "modern" HIAs. The air pollution data that are the basis of the study used for the CBA are from 1966-1987, or close to 30 years old! We question whether ERFs based on results from another continent using air pollution data from 30 years ago are robust enough to use in a CBA designed to project results nearly 20 years into the future, a near 50 year extrapolation. There is also the concern about whether an ERF derived from high air pollution levels and with a different air pollution mix, is relevant to current air pollution levels in Europe today. Indeed, the air pollution data in California are dominated by photochemical smog.*

**Response:**

These comments principally relate to whether the estimated effects of PM on chronic bronchitis change according to concentrations of PM and concentrations of other pollutants; i.e. according to the nature and intensity of the pollution mixture. We have no direct information about this, with regard to chronic bronchitis. However, where there is a lot of evidence, for example on the association of PM with 'acute' mortality and hospital admissions from time series studies, there is evidence of variation over time and location, but in general that variation is not explained by changes in the mixture within which PM occurs.

**Concern:**

*Likely, this ERF drastically over-estimates effects of low levels of PM alone.*

**Response:**

We do not agree with this opinion, which is entirely speculative.

- There are reasons why the effect may be *under*-estimated. Specifically, as noted in Vol 2, Section 8.3.3, the logistic regression analyses of Abbey et al (1995) considered new cases of chronic bronchitis 1977-87/88, in relation to estimated lifetime exposure. These analyses

“adjusted for covariates such as age, gender, education; *and* for respiratory symptoms in 1977 – this may imply over-adjustment for the effects of PM” (because it is likely that symptoms in 1977 are also in part a consequence of earlier exposure to PM).

- The estimated coefficient is high compared to many other studies of other health endpoints, but increasingly there is evidence that estimated coefficients are high when exposure estimation is based on fine spatial resolution, as it is here.

**Concern:**

*In fact, whether or not a threshold exists for this endpoint, and whether or not the ERF is specific to particulate matter, photochemical pollution, other gases present in ambient air, or a combination of these, has not been evaluated.*

**Response:**

We have used the same assumptions about threshold, or not, as for mortality from long-term exposure. This again assumes a similarity which we cannot prove but which seems to be a reasonable 1<sup>st</sup> assumption.

**Concern:**

***We are concerned about using a baseline disease rate from a single U.S.-based study and extrapolation to Europe without any adjustment or consideration of the uncertainties involved. Only limited information is provided on how baseline rates in the U.S. compare with those in Europe. It is well known that smoking is by far the major contributor to the production of bronchitis. One might guess therefore that the incidence of this disease might differ in Europe versus the U.S. according to difference in smoking rates. Nonetheless, we question the accuracy of basing a benefits analysis for bronchitis based on baseline disease rates from the U.S.***

**Response:**

We used background incidence rates from the AHSMOG study but only after cross-referencing them against two sources of information from Europe. As described in Section 8.3.2 of Vol 2.

- Results from some individual European studies report higher incidence rates
- Estimates from the Global Burden of Disease suggest lower rates
- From the AHSMOG study, we chose the lower of the two main candidate values

Note that the incidence rate of bronchitis in the AHSMOG study may be *low* because the smoking habits of the study group are much less than of the general population and, as noted above, smoking is the major risk factor for bronchitis. On the other hand, and again as noted in Volume 2, the AHSMOG authors' definition of chronic bronchitis is a relatively mild one, and this may lead to higher estimates of background incidence rates.

Considering all these factors, we think that the choice we made, from the AHSMOG study, is a reasonable one. It is not accurate to say that this transfer was done without consideration of the uncertainties involved.

**Concern:**

***We are concerned about using a valuation figure for bronchitis that is based on a single U.S. study (Krupnick et al.) and extrapolation to Europe without any adjustment or consideration of the uncertainties involved. Again, this figure comes from a single U.S.-based study. Costs of treating respiratory diseases in the U.S. markedly differ from those in Europe. For most medical related costs, figures in the U.S. are higher than those in Europe. Therefore, we question the accuracy of using the result from a single U.S. study for the CAFE CBA. Again, this decision appears to be driven by the comments of the reviewer, who strongly advocated using the results of his own single study. However, it is not clear why the authors of the CBA accepted this advice, seemingly without question.***

**Response:**

Volume 2 documents that the valuation of the Chronic Bronchitis end-point was based on the results of two studies undertaken in the US: (Viscusi et al, 1991; Krupnick and Cropper. 1992). Both studies were selected prior to the peer reviewers being selected. Indeed, it is worth emphasising that since there were no other studies available, there was little choice than to use these studies as a basis for an indicative range. Both studies used a Contingency Value (CV) approach and therefore did not refer to or include treatment costs. However, the UNICE comment correctly recognises – as does the text in

Volume 2 – that benefit transfer of this type incorporates uncertainty. One way in which this uncertainty is addressed is to use the results of the trade-off between the risk of chronic bronchitis and the risk of car accident fatality (Viscusi et al., 1991), in conjunction with a European value for accident risk, to derive part of the recommended unit value range, thereby avoiding sole reliance on US WTP values.

## **RESTRICTED ACTIVITY DAYS (RADS) AND MINOR RESTRICTED ACTIVITY DAYS (MRADS)**

### **Response:**

There are some limitations to the studies of RADs and MRADs that we have flagged up in the past; several of these – and others – are raised by UNICE. However, we have in the past quantified these effects and we consider that it makes sense to do so again. Indeed, for CAFE CBA we have found more supporting evidence than previous, though the Canadian study of Stieb et al (2002), and through background rates both in Canada and in the UK.

### **Concern:**

*We are concerned that assessment of these endpoints are based on the results of a single study, the Health Interview Study, as reported by Ostro et al., 1987, and Ostro and Rothchild, 1989.*

### **Response:**

In the words of the reviewer (Bart Ostro, who analysed these data): "First, the series of papers reporting this association are not really from "one study" but use six years of annual data collected by the National Center for Health Statistics". The concentration-response functions (CRFs) used were derived from meta-analyses of these six years. These were large studies – about 12,000 subjects per year, for the study of RADs, and from throughout the USA. We think that the results are robust enough to use in quantification.

### **Concern:**

*The ERFs are derived from a study from another continent and during the period of 1976-1981, or close to 30 years ago, when the air pollution levels were higher. We question the validity of extrapolating from results in 1976 to 2020 and beyond, a 40-year extrapolation. In the case of particulate matter, we question the tenuous exposure metrics used in this study.  $PM_{10}$  and  $PM_{2.5}$  levels were not measured as part of this study. Rather,  $PM_{2.5}$  levels were estimated from visibility data from airports. Results of other CBA assessments have indicated concerns about extrapolating results from high pollution levels to lower levels and resulting inflation of the ERF at higher levels.*

### **Response:**

There are limitations in the assessment of pollution, and also issues of transferability/ extrapolation to be considered. However, regarding transferability, estimates of the effects of PM seem to be generally quite robust with regard to pollution levels and the pollution mixture. The underlying study was US-wide, implying that results were based on multiple mixtures of pollutants. This may help with transferability.

HIA for some environmental pollutants, e.g. PAH, benzene and heavy metals, requires extrapolation from very high occupational concentrations to much lower environmental ones. Fortunately, the extrapolation needed for CAFE-CBA is not nearly as extreme as this.

### **Concern:**

*Further, there has been no assessment of whether RADs or MRADs would even be triggered by lower air pollution levels. In other words, the issue of threshold has not been explored at all for these morbidity endpoints.*

### **Response:**

When WHO considered the issue of threshold of effects it referred to both mortality and morbidity; its conclusions regarding absence of evidence for a threshold applies to both. Hence for the quantification and evaluation of health benefits, and in the absence of direct evidence for each endpoint, we consider that the WHO position for mortality is the best working assumption for RADs and MRADs also; i.e. quantify effects of *all* anthropogenic PM and effects of ozone above a cut-point of 35ppb (max 8-hr daily mean). Put differently, if pollution at low concentrations can lead to earlier mortality, then it seems reasonable to assume that it may also lead to health-related restrictions in activity.

**Concern:**

*In the case of particulate matter, we question the adjustment of an ERF for PM<sub>10</sub> to one based on PM<sub>2.5</sub> based on the simple mean ratio of these particles in urban air. The authors offer no biological explanation as to why such an adjustment is appropriate, or why fine PM would be expected to exhibit the same potency as coarse particles. Fine and coarse particles distribute differentially in the respiratory tract and as stated by the WHO and others, produce a different and separate spectrum of health effects. As described above, certain respiratory symptoms would be expected to be exacerbated more by exposure to coarse rather than fine PM, a finding consistent with the actual study results reported by Abbey et al., where stronger associations were observed for TSP than for PM<sub>10</sub> or PM<sub>2.5</sub> surrogates. It is not clear why the authors of the CBA choose to attribute all RAD related effects to fine PM.*

**Response:**

We have used the CRFs as in the published papers, in the metric of PM<sub>2.5</sub>, and linked this with the modelled pollution data – also in the metric of PM<sub>2.5</sub> – from RAINS. We do not mean to imply that there are no effects of coarse particles; rather, we are unable to quantify them. (This may imply some under-estimation of the total benefits.)

**Concern:**

*In the case of ozone, we question the plausibility of the association with MRAD, the ERF selected and how it is applied in the CAFE CBA. Ozone is a respiratory toxicant. In the study by Abbey et al., no association was reported between exposure to ozone and respiratory restricted activity days (RRADs). This begs the question, if those in the study were not restricted due to respiratory-related reasons, what biologically related reason accounts for their restricted activity that could be due to ozone exposure?*

**Response:**

On grounds of biological plausibility, we also would have expected an effect on respiratory RADs – that is why, in Vol 2, we drew attention to its absence. A possible factor is that there are far fewer RRADs than MRADs, and so less power to find a relationship.

**Concern:**

*Using a multi-pollutant model applied to air pollution data between 1976 and 1981, the author's report a positive association between 2-week average 1-hour ozone concentrations and MRAD.*

**Response:**

Rather than simply skip over the phrase 'using a multi-pollutant model', it may be helpful to remark that these studies, of RADs and MRADs, give coefficients for PM adjusted for ozone and for ozone adjusted for PM – something which elsewhere (see Attachment 2) UNICE highlights as desirable. (We agree.)

**Concern:**

*However, temperature is incorporated linearly in their model and is highly correlated with ozone, which decreases the certainty that ozone alone is causing MRADs.*

**Response:**

It is usual in these studies to model temperature in a linear way. I agree that in principle, there is the danger that possible confounding factors (such as temperature, here) have not been taken into account sufficiently well. Where this issue has been investigated carefully, e.g. in the work of Kalkstein and others using 'synoptic' weather variables, the pollution effect has remained robust to different ways of expressing climate. I would expect a good statistical analyst like Ostro to be aware of the possibility of residual confounding; but yes, this cannot be ruled out entirely.

**Concern:**

*In addition, there was high variance in the regression coefficients across the six years examined, with negative coefficients observed in 1977 and 1981 and a non-significant coefficient reported in 1976.*

**Response:**

Yes, there were some substantial year-by-year variations, but the overall estimated effects were similar when the extreme values were dropped from the meta-analysis.

**Concern:**

*This brings into question the conversion used by the authors of the CBA to convert an ERF based on 1-hr maximum levels to daily 8-hr averages. There is absolutely no question that higher peak concentrations of ozone produce more pulmonary effects than lower average levels. It is entirely possible that at lower average ozone levels, no respiratory effects and no MRADs would occur. However, the authors of the CBA choose to ignore this possibility and instead make the conversion to 8-hour average values, and extrapolate down to 35 ppb ozone, a level producing no clinical effects.*

**Response:**

These are two different ways of representing the same daily ozone pollution. Clearly, one method focuses on a 1-hr period, the other on an 8-hr period. Note that the pollution experienced by the individuals is unaltered; i.e. the individuals will have experienced exactly the same peaks and averages, regardless of how the pollution was measured; i.e. this is not about having the same amount of ozone 'administered' as a 1-hr peak or as a much lower average, over an 8-hour period.

What is at issue, then, is which metric is more appropriate, biologically, for estimating the health effects of daily variations in ozone. In the past, many studies used 1-hr peak ozone, because it resembled most closely the measures used in clinical studies of the effect of ozone – see later comments on clinical and epidemiological studies. More recently, thinking has moved to the viewpoint that ozone measured over a somewhat longer period – 6-hr or 8-hr daily max – may be a better metric for establishing effects on health. It is likely however that the choice of metric is not a big issue here.

**Concern:**

*We are concerned that all effects of air pollution on RADs and MRADs are being arbitrarily attributed to fine PM and ozone, with potential effects of other pollutants ignored. For example, we note that in a multi-pollutant model, the hypothesized effect of exposure to PM<sub>2.5</sub> on RAD did not persist following adjustment for carbon monoxide (Stieb et al, 2002).*

**Response:**

Some differences do occur with multi-pollutant models in different times and locations. As noted in Vol 2, Stieb et al. suggest that a reason why effects of CO rather than PM may persist in multi-pollutant models is because CO is measured more reliably – every day, compared with every 6<sup>th</sup> day for PM. (Measurement error in an exposure usually weakens the chances of finding a relationship, and the size of the estimated relationship.) The evidence on air pollution overall points most strongly to particles and to ozone as being the responsible pollutants. However, given the questions about aspects of the US studies, what we consider remarkable is not that there were some differences in results between US and Canada, but that there were so many similarities.

**Concern:**

*The health endpoints of RAD and MRAD are highly subject to socioeconomic confounding. In the study used to derive the ERFs, significant city-to-city differences in RAD rates were observed. This was likely due to socioeconomic factors and other factors that were not adequately controlled in the selected study. Some of these factors include time spent outdoors, building construction, health practices including how such days are recorded, age of the population, sex, race, education, income, marital status, temperature, employment conditions and rates, smoking rates, and many other factors. Even greater differences would be expected when considering cities in the U.S. versus those in Europe. Further, many of the socioeconomic factors that need to be controlled to identify the potential effect of air pollution are likely much more important than air pollution itself in the production of RADs and MRADS.*

**Response:**

Given the design, it was necessary to adjust for these factors, and that was done. If there remains a question, it is whether the adjustment done was sufficient. The experience of the ACS mortality study, where such issues were investigated in detail, is that whereas many factors are related to health and mortality more strongly than air pollution is, the estimated effect of air pollution is generally insensitive to these factors; i.e. the estimated effect of pollution is similar, regardless of how detailed is the adjustment for socio-economic and associated factors.

**Concern:**

*We are concerned that the authors use a RAD background rate taken from a U.S.-based study (ORNL/RFF, 1994). Socioeconomic factors such as disability rates, income status, unemployment rates, and various definitions of RAD will influence the background rates.*

**Response:**

Again, we agree in principle. In practice, however, and as noted in Volume 2, very similar background rates were found in the US study and in the Canadian one. The one snapshot we have of similar issues in Europe – derived from the UK General Household Survey – suggested substantially *higher* background rates in the UK.

Furthermore, in the core analyses impacts are estimated only for people at ages 15-64, whereas of course older people also have days when their activities are restricted for reasons of ill-health. In addition, for MRADS both the CRF and the background rates were based on people aged 15-64 *who were in employment*, and these would be expected to be more healthy – and so have fewer health-restricted days – than the general population. These considerations suggest that the effects may well be *under-* rather than *over-*estimated.

**Concern:**

*All of the above indicate significant concerns for the transferability of these ERFs for use in assessing RAD and MRAD in "average Europe", without any consideration for all of the uncertainties involved. Certainly, such "benefits" should not be included in the core CAFE CBA assessment, either for particulate matter or for ozone. In the case of particulate matter, any estimates that are made should be attributed to coarse PM rather than fine PM.*

**Response:**

While we recognise and accept that there are uncertainties, we do not accept that quantification of these endpoints should be restricted to sensitivity analyses only.

**LOWER RESPIRATORY SYMPTOMS (LRS) (INCLUDING COUGH) AMONG CHILDREN - PM**

**Concern:**

*Again, we question why the entire impact of particular matter on lower respiratory symptoms, defined as breathing disorders including wheezing, chest tightness, shortness of breath, and possible cough, are attributed to the fine particle fraction through converting an ERF based on PM<sub>10</sub> to one based on PM<sub>2.5</sub>. The authors provide no biological justification for this approach. It is entirely possible and plausible that coarse size particles landing in the upper respiratory tract contribute as much or more to these symptoms than do fine particles. Indeed, in the single study evaluating potential effects of both fine and coarse particles, essentially the same odds ratio was reported for both sizes of PM (1.091 and 1.086 for fine and coarse PM, respectively).*

**Response:**

The concentration-response function (CRF) given in Volume 2 is in the metric of PM<sub>10</sub>. In the absence of modelled PM<sub>10</sub> data being available, the conversion is a way of estimating relevant PM<sub>10</sub> concentrations, for linkage with the CRF. As discussed earlier for chronic bronchitis, linkage is not to be understood as implying that only fine particles have an effect on LRS.

**Concern:**

*More importantly, we question why this endpoint is being monetized at all. In the meta analysis conducted by WHO, the pooled result of 28 estimates of cough and particles in symptomatic children was close to 1.0 with a confidence limit of 95% indicating non statistical significance (WHO, 2004). In other parts of the CBA, the authors place high confidence and weight on the results of this meta analysis. It is not clear why in this case, they choose to ignore the results of the WHO meta analysis, and select a single study reporting a positive risk ratio for purposes of monetization.*

**Response:**

Our reasons are given in Vol 2, Section 10.4.2. The WHO meta-analysis is dominated by results from the 27 individual locations of the PEACE study; and, as noted by the PEACE study authors, by WHO, and in CAFE CBA Vol 2, there are methodological reasons (short time span, sometimes as little as six

weeks; an influenza epidemic) which cast doubt on the reliability of the PEACE estimates of the effect of air pollution on symptoms in the children – results which are generally negative. The 'single study' used to derive a CRF is (as described in Vol 2) also a meta-analysis, based on 12 studies (for cough) and 16 studies (for LRS), of which PEACE was treated as a single study.

## LOWER RESPIRATORY SYMPTOMS IN CHILDREN - OZONE

### Concern:

*For this analysis, the authors of the CBA again rely on the results of a single study that evaluated 91 children (Declercq and Macquet, 2000). We note that the risk ratios referenced in this study are higher than those reported for asthmatic children (Just et al., 2002). We doubt the assertion provided by the authors of the CBA that "there is no evidence that relative risks are higher among people with chronic respiratory disease than among the general population." Once again, the authors of the CBA rely solely on results of ecological epidemiology studies to make important conclusions.*

### Response:

Epidemiology and clinical studies give a different kind of evidence. Both are relevant to overall assessment of the evidence regarding the health effects of air pollution. It is well-established, however, that quantification is based on epidemiology, insofar as relevant studies are available. For the more severe health endpoints (mortality, hospital admissions, physician visits, days of restricted activity, medication usage), epidemiology is the only available source; clinical studies are impossible (neither ethical nor otherwise practicable). We base quantification of symptoms on epidemiology rather than on clinical studies because epidemiology approximates better to the conditions under which benefits do (or do not) occur in practice.

### Concern:

*The large human clinical database for ozone clearly supports a distinction in susceptibility between healthy individuals versus those with respiratory disease. Overall, in our view, the high results from this single study should not be used in the core analysis of the CBA.*

### Response:

We agree that we should make clear that this quantification is based on a small study only. If the relative risk as estimated from the CRF had been large, we would have recommended sensitivity analyses only, as we did for ozone effects on medication usage in children with asthma.

### References

*Abbey DE, Hwang BL, Burchette RJ, Vancuren T, Mills PK. (1995a). Estimated long-term ambient concentrations of PM10 and development of respiratory symptoms in a non-smoking population. Arch Env Health; 50: 139-152.*

*Abbey DE, Ostro BE, Petersen F, Burchette RJ (1995b). Chronic respiratory symptoms associated with estimated long-term ambient concentrations of fine particulates less than 2.5 micron in aerodynamic diameter (PM2.5) and other air pollutants. Journal of Exposure Analysis and Environmental Epidemiology; 5 (2): 137-159*

*Declercq C, Macquet V (2000). Short-term effects of ozone on respiratory health of children in Armentieres, North of France. Rev Epidemiol Sante Publique; 48 Suppl 2: 2S37-43.*

*Just J, Segala C, Sahraoui F, Priol G, Grimfeld A, Neukirch F. (2002). Short-term health effects of particulate and photochemical air pollution in asthmatic children. European Respiratory Journal, 20: 899-906.*

*Krupnick, A., Alberini, A., Cropper m; Simon N., O'Brien B., Goeree R., and M. Heintzelman (2002), "Age, Health, and the Willingness to Pay for Mortality Risk Reductions: A Contingent Valuation Survey of Ontario Residents," Journal of Risk and Uncertainty, 24, 161-186.*

*ORNL/ RFF (1994). Estimating Externalities of the Coal Fuel Cycle. report 3 on the external costs and benefits of fuel cycles. A study by the US Department of Energy and the Commission of the European Communities. Prepared by Oak Ridge National*

*Laboratory and Resources for the Future. McGraw Hill.*

Ostro B.D. (1987). *Air pollution and morbidity revisited: A specification test. J Environ Econ Manage* 14, 87-98.

Ostro B.D. and Rothschild S. (1989). *Air pollution and acute respiratory morbidity: An observational study of multiple pollutants. Environ Res* 50, 238-247.

Stieb DM, Smith-Doiron M, Brook JR, Burnett RT, Dann T, Mamedov A, Chen Y. (2002). *Air pollution and disability days in Toronto: Results from the National Population Health Survey.*

Tiitanen P, Timonen KL, Ruskanen JJ, Mirme A, Pekkanen J (1999). *Fine particulate air pollution, resuspended road dust and respiratory health among symptomatic children. Eur Respir J*; 13: 266-273.

WHO (2004); prepared by Anderson HR, Atkinson RW, Peacock JL, Marston L, Konstantinou K. *Meta-analysis of time-series studies and panel studies of particulate matter (PM) and ozone (O<sub>3</sub>). Report of a WHO task group. World Health Organisation. (<http://www.euro.who.int/document/e82792.pdf>; accessed November 2004.*

## Attachment 2

### DETAILED COMMENTS ON USE OF EXPOSURE RESPONSE FUNCTIONS IN THE FINAL CAFE CBA METHODOLOGY

#### Response:

The CAFE CBA's team response to UNICE's concerns about quantification of mortality is briefer than that for morbidity, for two reasons:

- a. These issues have been discussed previously; and
- b. As explained in Methodology, Vol 2, the main decisions are based on recommendations by the Task Force on Health (TFH) of the WHO-UNECE Convention on Long-Range Transboundary Air Pollution (CLRTAP). These were implemented in RAINS and adopted for the CBA analysis because (i) of the status of WHO recommendations and (ii) for consistency of the overall CAFE process. However, we recognise that these are important issues and we trust that the present response will show that the decisions taken were well-informed and fair.

## CHRONIC PM MORTALITY

### Concentration Response Functions

#### Concern:

In our previous comments, in order to avoid multiple counting of mortality impacts, we suggested using concentration response functions (CRFs) that are adjusted for effects of other pollutants proposed for valuation. In addition, we suggested that careful consideration should be given in cases where European specific CRFs are not available and the methodology calls for "transferring" functions from other regions or continents, e.g. from the U.S. to Europe.

For the critical health effect of chronic PM mortality, the authors of the CBA have directly transferred the value of 6% change in mortality per 10 ug/m<sup>3</sup> of PM<sub>2.5</sub>, as derived from the American Cancer Society (ACS) Study for use in the CAFE CBA. While this proposal follows the general recent WHO-UNECE advice, we believe it is unrealistic to represent the results of this complex study with a single risk estimate.

The results of the thorough reanalysis of this study by Krewski et al. (2000, 2003) clearly demonstrate effect modification by education and other factors such as temperature variation and population change, attenuation of particle effect when spatial correlation was considered, and most importantly, strong attenuation of the particle effect when sulphur dioxide was simultaneously considered in the model. Thus, for example, when sulphur dioxide was simultaneously considered in the model for all cause mortality, the coefficient for sulphates was reduced from 0.17 to 0.05 and that for fine particles from 0.18 to 0.03, or reductions of 71% and 83%, respectively. Both values adjusted for sulphur dioxide were also statistically insignificant (see HEI report, summary table 6, page 31, or table 37 page 184).

We note that the levels of sulphur dioxide that exist in Europe today are much lower than those present in the U.S. over the period when the ACS study was conducted. We also note that human clinical and toxicology studies clearly demonstrate that even low concentrations of SO<sub>2</sub> and acidity significantly enhances the toxicity of particles, i.e., producing a reducing-type air pollution (Amdur, 1991). Finally, we note that very recent studies on short-term effects of PM<sub>10</sub> on cardiovascular diseases in eight European cities clearly showed that impacts of particles in European cities were half those found in cities in the United States using the exact same study design and protocol (Le Tertre et al, 2002). Hence, careful thought must be given to transferring CFRs from studies in the U.S to Europe.

Whether ambient exposure to sulphur dioxide produces an independent risk for mortality, as suggested by Krewski et al., acts as a surrogate for other pollutants in the air pollution mix, as suggested by the WHO, or actually increases the risk of PM, as suggested by the clinical and toxicology studies, is arguable. However, using the 6% value, alone, without adjustment, does not provide a fair and accurate way of estimating the risks in Europe under the CAFE program. In response to this concern, we note that the draft CBA included a proposal to conduct a sensitivity analysis using a coefficient of 4%. This figure has been widely used, for example, by WHO in previous Global Burden of Disease analyses. It is also reasonable to consider this value in relation to the likely long time lag for chronic effects, making the

1979-83 time period a more relevant period to consider from a biological perspective. Therefore, we support the concept of using an alternate to the 6% figure as proposed in the previous draft CBA as part of an uncertainty analysis. However, in addition to the 4% figure, we suggest using a coefficient of 1% in addition to 4%, based on data by Krewski et al. which provide a more complete adjustment for removing the effects of SO<sub>2</sub> and transferring the CRF for use in the CAFE CBA. One could consider a higher figure such 9%/10 ug/m<sup>3</sup> as reported from in the 6 cities study. However, in this case, one should also consider a no chronic mortality scenario reported in the scientifically valid VA study by Lipfert et. Al. The no chronic mortality scenario as well as figures lower than 6% were considered relevant by a recent expert review group convened by USEPA. In fact, the CAFE program should strongly consider use of a risk distribution function ranging from 0 to 6-9%.

**Response:**

We think that there are differences of interpretation between WHO-UNECE and UNICE on aspects of the ACS study but no substantial difference of opinion on what to do.

- We agree that different statistical models from the many analyses of the ACS study data give rise to different coefficients. These differences, and the possible reasons for them – e.g. what may be the role of SO<sub>2</sub> – were taken into consideration by WHO-UNECE Task Force on Health when it recommended that a 6% increase in hazard rates (age-specific death rates) per 10µg/m<sup>3</sup> PM<sub>2.5</sub> as the preferred best estimate, to be used as the core value for quantification.
- We note that UNICE accepts (without necessarily agreeing with) use of this estimate for the core analyses
- We also agree – as does TFH of WHO-UNECE – that variations around this estimate should be considered in the uncertainty analyses. This is being done; results will be reported in Volume 3 of the CAFE CBA series, which is currently under preparation. What we have done is to use Monte Carlo methods to take account of sampling variation (i.e. the size of the confidence interval) around the preferred central estimate, as part of an integrated analysis of several sources of uncertainty. This approach encompasses the range of values proposed for uncertainty analysis by UNICE, and also includes some higher values to reflect the findings of other recent studies – noted in Vol 2 but not mentioned by UNICE – which suggest that exposure estimation at a greater level of detail than was possible in the general ACS study may lead to higher estimates of relative risk.

**Thresholds/Cutpoints**

**Concerns:**

*In the final CBA, the authors take the general approach that all health effects produced by all ambient pollutants, including PM, occur at any level of exposure with a linear exposure response. To justify this approach, the CBA makes reference to the recent WHO review of the health effects of PM, ozone, and nitrogen dioxide.*

**Response:**

As noted by UNICE, this issue was considered in detail in the WHO review and, for ozone, in the supplementary WHO review giving answers to follow-up questions from CAFE. That evaluation did conclude that the evidence did not identify a threshold concentration below which ambient PM has no effect on health. WHO review also concluded that “it is likely that within any large human population there is such wide range of susceptibility that some subjects are at risk even at the lowest end of the concentration range”. The rationale to the above answer explains, that: “Effects on mortality and morbidity have been observed in many studies conducted at exposure levels of current interest. If there is a threshold, it is within the lower band of currently observed PM concentrations in Europe. As PM concentrations are unlikely to be dramatically reduced in the next decade, the issue of the existence of a threshold is currently of more theoretical than practical relevance.”

The CAFE CBA has adopted this conclusion of the WHO review.

**Concern:**

*The WHO asserts that with any large population, there may be some susceptible individuals that react at some low level. Therefore, it may not be possible to observe thresholds at the population level. This construct is theoretical and as such cannot be proven or disproven. However, the long-standing assumption in human health risk assessment is that population susceptibility is not "infinite" and there is a threshold (i.e., some dose below which there is no adverse effect) for non-cancer health effects.*

The unequivocal no threshold approach also runs contrary to, for example, a very recent assessment by the USEPA, wherein the agency employs the approach of extrapolating no further than the **lowest measured level (LML)** from the study used in the risk analysis (EPA, 2005). These values are 7.5, 10, and 11  $\mu\text{g}/\text{m}^3$  for the ACS-extended, ACS Krewski, and 6-cities study, respectively (EPA, 2005). As discussed by EPA, the greatest uncertainty in PM models is whether or not there is a threshold and at what level of exposure (OAQPS, 2003, EPA, 2005). Furthermore, the approach used in the CAFE CBA ignores evidence within the PM literature that the CR functions may not be linear and that a threshold model may be more appropriate.

**Response:**

The qualitative guidance from the WHO reviews was considered further by TFH of WHO-UNECE with a view to quantification. TFH recommended that:

- Effects of PM on health and mortality be quantification as the effects of *anthropogenic* PM only, but without threshold; and
- Effects of ozone on health be quantified only above a cut-point of 35ppb ozone ( $70 \mu\text{g}/\text{m}^3$ ) in the core analyses, with quantification of effects down to zero to be included in sensitivity analyses only.

The CAFE CBA team has adopted this approach. Note that an 'unequivocal no threshold position' would have led to quantification, in the core analyses, of effects of PM from all sources (i.e. not just anthropogenic) and of ozone down to zero, i.e. with no cut-point. In reaching these conclusions, TFH considered carefully the reliability of extrapolating the ACS study results downwards, somewhat below the lowest observed level of the ACS study; and concluded that the limited extrapolation implied in quantifying anthropogenic PM without threshold was likely to lead to better estimates than would be got by limiting the analysis at the lower end of the ACS-extended study from which the 6% core estimate was derived.

It has to be kept in mind that the EMEP/RAINS model systematically underestimates the levels of PM, partly due to the fact that non-anthropogenic sources are not covered and partly due to the lack of secondary organic aerosol, of which a substantial part may be of anthropogenic origin. The total  $\text{PM}_{2.5}$  mass estimated by the EMEP/RAINS model only account for about 80 percent of the observed  $\text{PM}_{2.5}$  mass. The effects of  $\text{PM}_{2.5}$  would therefore be systematically underestimated in that respect.

**Concern:**

*In our previous comments, we expressed concern for the approach of assessing chronic  $\text{PM}_{2.5}$  mortality with no threshold or cutpoint. In our comments, we provided the various lines of scientific evidence supporting consideration of a threshold. These arguments are provided again briefly below.*

**Response:**

We understand that the issue is considered as important by UNICE. However, as UNICE has noted, the issues have been raised and considered previously.

In the absence of new information or insights, we will not rehearse once again the arguments which led WHO to its conclusions, which we then accepted for CAFE. We think it unlikely that doing so will cause UNICE to change its position, just as the statement of position proposed by UNICE (presented and discussed by WHO working groups before their final conclusion was reached) – and re-iterated below – was not convincing to the various WHO expert groups.

**Concerns:**

**1) Thresholds or no linear results have been reported in time series studies**

*It has been suggested by WHO and others that the association between PM and mortality increases linearly with increasing concentrations, with no evidence of a threshold. (Pope et al., 1992; Schwartz, 1994; Dockery et al., 1994). This conclusion is subject to question for several reasons, including:*

- *No systematic attempt has been made to explore non-linear relationships and the shapes of C-R functions have not been adequately investigated as to whether a linear relationship can be assumed. An assumption of linearity is not valid until careful investigation is conducted considering the effect of co-pollutants, assessment of statistical significance, and selection of best-fitting models.*

- An assumption of linearity means that the risk estimate is usually driven by the highest risk (Moolgavkar and Luebeck, 1996).
- The linear, no-threshold models within the PM time-series literature usually contained no co-pollutants. Moolgavkar and Luebeck (1996) pointed out the significance of this exclusion using data from Philadelphia.
- Thresholds and non-linear exposure response relationships have been reported in a number of studies.

Schwartz and Zanobetti (2000) suggest that a finding of no threshold is not unexpected, as human populations are diverse with widely varying risks, so that C-R functions should resemble an S-shaped curve with no threshold. Exposures and risks are low so the data are at the low-end of the curve where the linear term dominates. This hypothesis is said to be supported by empirical evidence, such as a minority of smokers get bronchitis after a lifetime of smoking while a few smokers get bronchitis after a few years of smoking. However, there is also evidence that does not support this idea, which is ignored by WHO and the authors of the CBA methodology. For example, Gori and Mantel (1991) estimate a threshold of 4-5 cigarettes/day for increased risk of lung cancer, CHD, or respiratory disease mortality. This area below the smoking threshold is in the same exposure range as ambient PM concentrations. Therefore, an alternative interpretation is that the apparent linearity at the low-exposure portion of the C-R model is due to statistical noise, rather than the absence of a threshold. This is biologically plausible and is supported by empirical evidence when the shape of the curve is more fully explored.

Daniels et al. (2000) attempted to test the correctness of the linear assumption in 20 cities from NMMAPS. The posterior modes from threshold models were 15  $\mu\text{g}/\text{m}^3$ , 15  $\mu\text{g}/\text{m}^3$  and 50  $\mu\text{g}/\text{m}^3$  for total, cardiovascular/respiratory (CVDRESP), and other-cause mortality, respectively. Using only AIC criteria, the linear model was superior to a non-threshold approach except for other cause mortality. In the subsequent revised analysis of these 20 cities, the results were similar (Daniels et al., 2004). However, the HRC was not convinced that AIC was the appropriate criterion on which to choose the appropriate model and other tests show variable results. A focused significance test, for example, rejects the log linear model.

The Health Review Committee of HEI (Daniels et al., 2004) also noted that because concentration-response functions for "Other Cause" mortality were non-linear, then concentration-response functions for Total mortality must also be non-linear as well, even if CVDRESP C-R functions appear to be linear. Furthermore, the HRC (Daniels et al., 2004) was less convinced than Daniels et al. (2000) of linearity based on AIC criteria, as AIC is biased in favour of a log linear model. The notion of comparing models to assess "scientific theories" (shape of the C-R function) has not been tested. In sum, the HRC "has some reservations" regarding the investigators' conclusion that the log linear no-threshold model is preferred.

Other authors have reported non-linear results for five of these 20 cities. These are Los Angeles, Chicago, Phoenix, New York, and Philadelphia (Moolgavkar, 2003; Smith and Rothman, 2000). Furthermore, other authors have found non-linear relationships for a number of cities, including St Louis (Samet et al., 1995), Utah Valley (Samet et al., 1995), Birmingham (Smith et al., 1997, 2000), Phoenix (Smith et al., 2000), Philadelphia (Samet et al., 1995; Nicolich and Gamble, 1999), Chicago (Smith, 1997; Moolgavkar, 2003; Roberts, 2004), and Los Angeles (Moolgavkar, 2003). The draft CBA methodology essentially ignores this published evidence of non-linearity within the time-series literature.

It should also be noted that incompletely controlled residual confounding and exposure measurement error tend to bias the CR function toward linearity. Since the statistical test used to assess linearity is not very powerful, the results can be influenced depending on whether one is testing to see if the curve is different from linear or different from non-linear.

Finally, it does not appear useful to test the linearity hypothesis based on combined cities, as was done by Daniels et al. (2000, 2004), given that it is not appropriate to statistically combine results from such heterogeneous data, at least without some consideration of possible regional and spatial effects.

**2) Thresholds, No Effect Levels, and Non Linear Responses Have In Fact Been Reported in Chronic Studies, Including the Study Proposed For Assessing Chronic PM Mortality in the Draft CBA.**

The most consistent finding from cohort studies appears to support a threshold above the current annual standards.

- Six Cities:

The conclusions that can be drawn from this study are severely limited, because only 6 data points are available, making it difficult to determine the shape of the curve. Within the limitations of these data, non-linearity has been mentioned as a possible C-R function (Krewski et al., 2000). Using a LOAEL/NOAEL approach, the only significantly increased risk of mortality in the six-cities study is a LOAEL of 29.6 ug/m<sup>3</sup> for males in Steubenville and a NOAEL of 20.8 ug/m<sup>3</sup> for women in Harriman.

- ACS:

The draft CBA methodology makes only passive mention of the fact that a threshold for certain forms of PM (e.g. sulphates) has been reported in the key study proposed for valuation of chronic PM mortality. However, the limitations associated with assuming linearity in the face of these non-linear C-R functions should be made more explicit. The C-R functions from the original ACS data appear to be non-linear (Krewski et al., 2000; Abrahamowicz et al., 2003). Therefore, nonparametric model results may take precedence over the linear Cox proportional-hazards models, given that the Cox model assumptions were not met (Krewski et al., 2000). Furthermore, the nonparametric models of the updated ACS cohort (Pope et al., 2002) indicate possible thresholds above 30 ug/m<sup>3</sup> PM<sub>2.5</sub> for most mortality estimates, except for cardiopulmonary mortality, which appears to have a threshold of ~20 ug/m<sup>3</sup>.

- AHSMOG (Abbey et al., 1999):

Threshold models showed significantly increased risk of mortality for males only at PM<sub>10</sub> concentrations >100 ug/m<sup>3</sup>, with no increased risks for women above this threshold. No significantly increased risk of total mortality was observed for either men or women using linear models.

- VA study (Lipfert et al., 2000):

No significant associations between PM<sub>2.5</sub> and mortality were reported, suggesting either no effect from PM or a threshold at some concentration above the maximum of 42 ug/m<sup>3</sup> PM<sub>2.5</sub>.

The draft CBA does discuss many of the above points. However, the authors essentially conclude that there is either no evidence for a threshold or that if there is a threshold, it is at a level below current PM regulations. This continued insistence that there is no health-effect threshold, in the face of evidence to the contrary, is disconcerting.

The published results of the cohort studies indicate there are inconsistencies in estimated risks in cohort studies and that the results are model-dependent. In fact, there is a somewhat consistent pattern of no significant effects at concentrations above the current annual standard, suggesting there may be no deleterious health effects at current ambient PM concentrations (see table). However, this evidence is not well presented at all by the authors of the draft CBA, who continue to conclude that the cohort studies provide clear and consistent evidence of PM health effects below current regulatory levels. The contrast in results between those presented in the draft CBA methodology and those in the table below highlights this discrepancy. The authors of the draft CBA methodology should explain why there is such wide variability in results, instead of trying to hide this inconsistency by presenting results in several formats in different tables.

**Table 1. RR of Total Mortality in Cohort Studies of PM**

Six Cities (* Krewski et al., 2000; #Villeneuve et al., 2002)	ACS reanalysis (Krewski et al., 2000; Abrahamowicz et al., 2003)	ACS update (Pope et al., 2002)	AHSMOG (Abbey et al., 1999)	VA (Lipfert et al., 2000)
* M: LOAEL = 29.6 F: NOAEL = 20.8 # Stratified by age <60: 1.21 (1.05-1.39) >60: 1.05 (0.96-1.13)	*Linear Cox Model <b>1.05 (0.85-1.30) /</b> range  # <b>Questionable</b> <b>accuracy of linear</b> <b>model</b>	*Linear Cox Model <b>1.06 (1.02-1.11)</b>  # <b>No significant</b> <b>effect in nonpara-</b> <b>metric smooth</b>	<b>Threshold Model:</b> <b>PM<sub>10</sub> &gt; 100 ug/m<sup>3</sup></b> <b>M: 1.12 (1.01-1.24)</b> <b>F: 0.94 (0.86-1.03)</b> <b>Linear Model</b> <b>M: 1.04(0.99-1.10)</b> <b>F: 0.97 (0.93-1.02)</b>	Linear Model <b>0.90 (0.85-0.95)</b>  Threshold > ~40 ug/m <sup>3</sup> PM <sub>2.5</sub>
*HRC: too few data points for regression # adjustments for changes in PM; effect modification by age	*Krewski, Regional Adj (Table 47)  #Abrahamowicz (2003)	Single pollutant model w/o SO <sub>2</sub> *Table 2 (Average 1979-2000) # Figure 2	Table 3, Threshold significant for males; Linear model results not statistically significant for M or F	from Table 8-II, EPA CD

RRs are for 10 ug/m<sup>3</sup> change in PM for linear models unless noted otherwise.

**3) Various types of anthropogenic PM demonstrate low toxicity and are therefore not unlikely to contribute to severe health effects such as mortality.**

As noted in the WHO report on the follow-up questions from CAFE and in many reviews such as by Schlesinger et al., (2003) there are many forms of anthropogenic PM that exhibit low toxicity, including sulphates, nitrates, ammonium salts, and chlorides. Other forms of non-anthropogenic PM such as sea salts and wind blown dust also exhibit low toxicity. These forms of PM are unlikely to contribute to severe health effects such as mortality. In total, these forms of PM comprise a fair high fraction of the ambient PM. These data support the likelihood of a threshold below which no mortality is expected.

**4) Human clinical and toxicology studies demonstrate no or only marginal effects at levels above ambient.**

A full review of this topic is beyond the scope of the CBA. However, a recent thorough review of this topic is presented in chapter 7 of the USEPA draft criteria document for particulate matter (EPA, 2004). In summary, as noted in this chapter, the toxicology data available today provide little basis for concluding that any specific components of PM have substantial respiratory or cardiac effects at current ambient levels of exposure (see pages 7-36, 7-85, and 7-206).

**5) Finally, all of the plausible yet unproven biological mechanisms for PM toxicity describe threshold-based modes of action.**

A full review of this topic is beyond the scope of the CBA. However, a recent thorough review of this topic is presented in chapter 7 of the USEPA draft criteria document for particulate matter (EPA, 2004). The plausible mechanisms of action include various cardiac and respiratory modes of action. However, all of these plausible yet unproven mechanisms of action involve phenomena that are expected to exhibit a biological threshold.

Therefore, we continue to recommend the use of a cutpoint for chronic PM<sub>2.5</sub> mortality in an uncertainty scenario.

**Response:**

Notwithstanding the case put persuasively by UNICE, above, the sense of the meeting of TFH of WHO-UNECE in Bonn in May 2004 was that there was insufficient evidence *in favour* of a threshold or cut-point for the effects of anthropogenic PM to justify the use of a cut-point even in sensitivity analyses.

**PM Metrics**

**Concern:**

*In our previous comments, we expressed concern that the entire burden of PM mortality was being placed on one size of particulate matter, PM<sub>2.5</sub>. Recently, the WHO has concluded that coarse PM, defined as PM<sub>2.5-10</sub> produces health effects independent of those from fine PM and therefore should be controlled (WHO, 2004a). In the recent WHO meta analysis of time series studies, a central European mortality estimate was provided for PM<sub>10</sub>; insufficient data were available to establish a similar estimate for PM<sub>2.5</sub> (WHO, 2004b). Finally, results of the EU-funded HEPMEAP and RAIAP programs clearly show that coarse PM produces more health effects (e.g. inflammation, asthma-related endpoints) than does fine PM (AIRNET, 2004; Schins et al., 2004). In our view, it is not appropriate to arbitrarily exclude coarse PM from the CBA mortality analysis. This will limit the possibility for policy measures directed at reducing coarse PM which may be more cost effective than some of those directed at fine PM. Therefore, we recommend including in the CBA a method to assess acute PM<sub>2.5-10</sub> mortality separate from any assessment of PM<sub>2.5</sub>, i.e., this should not be added to the assessment of PM<sub>2.5</sub>.*

**Response:**

CAFE CBA follows the WHO guidance by (i) quantifying mortality from long-term exposure to PM in terms of anthropogenic PM<sub>2.5</sub> and (ii) *not* adding to it any estimated effects of mortality from short-term exposure (daily variations), to avoid a danger of double-counting. We agree that there may be an effect of coarse PM on mortality; that the evidence for this is from short-term studies only (i.e. time series studies); and that by not including mortality effects of short-term studies we may be ignoring a valid component of the overall burden of mortality attributable to PM, via the effects of the coarse fraction.

For completeness, and to provide a better framework for discussion, we do however estimate the number of annual earlier deaths attributable to short-term exposure to PM, only some of which are believed to be due to the coarse fraction. Results showed that the estimated value is small in monetary terms, especially if (as UNICE recommends elsewhere in this letter) the VOLY rather than VOSL approach to mortality valuation is taken; and so the omission of coarse particle mortality effects is unlikely to have a substantial influence on policies, whether CBA or cost-effectiveness policies.

**PROPOSED METHODOLOGY FOR ASSESSING OZONE**

**Concentration Response Functions**

**Concern:**

*In our previous comments, we suggested that worst case CRFs should not be used systematically. Rather, CRFs selected for the primary analysis should be the best, most accurate and central ones. In particular, we referenced the results of the recent meta analysis of European time series studies which for ozone reported a mortality figures of 0.3% / 10 ug/m<sup>3</sup> ozone and 0.2% / 10 ug/m<sup>3</sup> after adjustment for publication bias, rather than the single city value of 0.059 (Sunyer et al., 1996) included in previous CBA evaluations. Therefore, we support the proposal in the CBA to use the results of the meta analysis.*

**Response:**

We chose to base results, where practicable, on the WHO meta-analysis, which in turn was based on European studies. The meta-analysis result agrees with the recently published estimate from APHEA2 study based on data from 23 European cities ((Gryparis et al, Am J Respir Crit Care Med, 170:1080-87, 2004). The earlier single-city value was intended to reflect the evidence at the time when it was selected, during 1997-98.

**Concern:**

*Since publication bias is a valid concern and has been adjusted for in the meta analysis, we suggest using the value of 0.2% as the best estimate for the primary analysis rather than 3%. We see no justification for claiming to consider publication bias and simultaneously not using the adjusted more accurate value. For example, trade-offs for lag selection biases, which can potentially decrease risk*

estimates as some assert for single day lags, or increase estimated risks, as is known in the case of selection of the highest result from amongst many lags, should be addressed separately from publication bias. We note that the protocol for the meta analysis likely leaned in the direction of increasing risk, since the lag selection ordering was as follows: 1) lag focused on by the author; 2) most statistically significant; 3) largest estimate. As a minimum, the 0.2% value should be considered as part of an uncertainty analysis. Use of the 2% value would also be more similar to the results of the high quality NMMAPS study, which coincidentally, the reviewer, Bart Ostro, recommended for consideration.

**Response:**

The issue was discussed at TFH in May 2005; we have adopted the TFH recommendation. In commenting on the NMMAPS study, Bart Ostro noted that "the NMMAPS study may underestimate the impact of mortality due to the modelling methodology used to control weather factors".

**Concern:**

*However, in our previous comments, we noted the general concern for using CRFs that are not adjusted for effects of other pollutants proposed for valuation that are presumed to cause the same health effect. We note that both 0.03% and 0.02% values noted above were derived using single pollutant models. Potential confounding with PM<sub>2.5</sub>, or for that matter any other pollutant, was not considered.*

**Response:**

We agree that adjustment for co-pollutants is better, but many studies do not do this, and the WHO meta-analysis is based on single-pollutant studies. However, whereas with many pollutants adjustment for one leads to lower estimated effects for another (because the pollutants co-vary positively), this is not necessarily the case with ozone and PM, as indicated in the recent article from APHEA2 study (Gryparis et al, Am J Respir Crit Care Med, 170:1080-87, 2004).

**Concern:**

*We also emphasize the general concern that for ozone, the levels observed on ambient monitors provide a very poor surrogate for personal and indoor exposure, with correlation's on the order of 20-30%. For example, data from Liu and Koutrakis (1993) and others indicate that outdoor ozone concentrations show substantial spatial variation and use of fixed-site measurements to estimate personal exposure can result in significant errors. For this reason, there are valid concerns for using results of ozone time series studies in a CBA, particularly when the CRFs proposed for use are not adjusted for PM<sub>2.5</sub>, a pollutant which does penetrate indoors and has a higher ambient to personnel correlation, and is presumed to cause the exact same health endpoint, i.e., acute mortality.*

**Response:**

We agree that (contrary to PM<sub>2.5</sub>) ozone at fixed-point monitors is not always a good indicator of personal exposure to ozone. However, the relationships from time series studies linking ozone (measured in urban background locations) with mortality, respiratory hospital admissions and other (respiratory) morbidity effects are widely regarded as causal, and usually included in HIAs/ CBAs.

**Concern:**

*For this very reason, the USEPA has not proposed to assess acute ozone mortality in their primary analysis of the cost and benefits of the clean air act (EPA, 2003).*

**Response:**

It is our understanding that:

- the USEPA deferred a decision on whether or not to include ozone and mortality in its proposals (2003) on quantifying the benefits of the US Clean Air Act;
- its reservations were based on different seasonal patterns of effects – an issue discussed by TFH when deciding to proceed with a quantification;
- it proposed to commission one or more meta-analyses, and review its plans once results of these became available.

**Concerns:**

*Therefore, based on the concerns described above, we strongly suggest conducting an uncertainty analysis based on the assumption for ozone occurs, or if it occurs, cannot be separated from the effects of PM.*

**Response:**

We do not think that the UNICE recommendation is warranted or workable. We do not agree that ozone effects cannot be distinguished from those of PM. To treat them as indistinguishable would imply omitting ozone and mortality, because the CAFE CBA methodology includes mortality effects of PM derived from cohort studies, and mortality effects of ozone derived from time-series studies.

**Thresholds/Cutpoints**

**Concerns:**

*In our previous comments, we expressed concern for the universal assumption that there are no thresholds for all health endpoints of all pollutants, and that a linear concentration response exists for all health endpoints. This dramatically inflates the benefits and does not reflect the current state-of-science. In the case of ozone, we noted various lines of evidence indicating a threshold exists for acute ozone mortality as reported in time series studies including;*

- 1. there is a clear seasonal pattern of effect with non-significant findings in winter when ozone is low (HEI, 2000);*
- 2. higher concentrations of ozone clearly produce more effects in human clinical and toxicology studies and as quoted in the recent EU-funded review under AIRNET, "Threshold of toxicity: there is ample documentation that ozone exerts its toxicity in a dose-dependent way" (AIRNET, 2004b);*
- 3. thresholds have in fact been reported in the majority of time-series studies (Walton, 2003) particularly those that adjusted for effects of other pollutants (e.g. Moolgavkar, 95);*
- 4. clear no effect levels for even minor effects in humans are reported in clinical studies, including studies in potentially susceptible groups such as asthmatics, and those with mild to moderate COPD and hypertension (see reference sub list 1);*
- 5. there remains no biological mechanism of action to explain acute ozone mortality and the existing and well documented mechanisms of actions are threshold-based;*
- 6. exposure misclassification in time series studies obscures finding thresholds, particularly for pollutants such as ozone for which the correlation between centrally monitored concentrations and personal exposure to ambient pollutants are weak (Brauer et al. 2000).*

*Therefore, we support the proposal in the revised CBA to include a cutpoint below which mortality will not be quantified. However, we suggest a value of 50-60 ppb rather than 35 ppb. A range of 50-60 ppb is consistent with the concerns expressed by WHO for increased uncertainty in extrapolating below 60 ppb (WHO, 2004). Specifically, in the detailed discussions, there was increased uncertainty in extrapolating below 60 and 40 ppb for mortality and morbidity, respectively.*

**Response:**

As noted earlier, CAFE CBA adopted the recommendation of TFH of WHO-UNECE. The TFH choice of cut-point was made by an expert group (Bonn, May 2004) which was well-informed both about the epidemiology and about ozone modelling/ background levels of ozone. The reviewer, Bart Ostro, was in favour of quantification down to background levels, presumed to be lower than 35ppb. The most recent evidence from APHEA supports either quantification to below 35ppb, or use of a higher risk estimate if a cut-point is used. The cut off level is not to acknowledge that there is a threshold for the effects, but acknowledges that the lower the levels of ozone the larger the uncertainties of the effects. The cut off level should therefore be viewed as level relevant for evaluating different policy options and their effectiveness of reducing the health impacts due to ozone.

**References**

*Abbey DE, Nishino N., et al. (1999) Long-term inhalable particles and other air pollutants related to mortality in non-smokers. Am J Respir Crit Care Med 159:373-382.*

*Abrahamowicz M, Schopflocher T., et al. (2003) Flexible modelling of exposure-response relationship between long-term average levels of particulate air pollution and mortality in the American Cancer Society study. J Toxicol Environ Health A 66(16-19):1625-54.*

*AIRNET (2004). RAIAP/HEPMEAP/AIRNET Workshop. Abstracts and Workshop Presentations, 11-12 March, Bilthoven, The Netherlands.*

AIRNET (2004b). *State of the Art in Toxicology of PM, Ozone, Nitrogen Dioxide, Volatile Organics, Including PAH, Carbon Monoxide, and Sulphur Dioxide.*

Amdur, M. *Air Pollutants.* In Casarret and Doull's *Toxicology the Basic Science of Poisons*, , 4th ed., 1991.

Brauer 2000. *Exposure misclassification and threshold concentrations in time series analyses or air pollution health effects.* *Risk analysis* 22: 1183-1193.

Daniels MJ, Dominici F., et al. (2000) *Estimating particulate matter-mortality dose-response curves and threshold levels: an analysis of daily time-series for the 20 largest US cities.* *Am J Epidemiol* 152(5):397-406.

Daniels MJ, Dominici F., et al. (2004) *The National Morbidity, Mortality, and Air Pollution Study, Part III: PM10 Concentration-Response Curves and Threshold for the 20 Largest US Cities.* Research Report 94. Boston, MA, Health Effects Institute.

Dockery DW and Pope CAI. (1994) *Acute respiratory effects of particulate air pollution.* *Annu Rev Public Health* 15:107-132.

Health Effects Institute (HEI) (2003)

EPA 2003. *Analytical blueprint for benefits and costs of the clean air act.* Pg. 6-13.

EPA (2004) *Air quality criteria for particulate matter. Fourth External Review Draft.* EPA/600/P-99/002bD.

EPA (2005). *Second Draft EPA PM Staff Paper, Docket No. OAR-2001-0017, FRL-7866-9.*

HEI (2000). *The National Morbidity, Mortality, and Air Pollution Study Part II. HEI Report 94, Part II.*

Gori GB and Mantel N. (1991) *Mainstream and environmental tobacco smoke.* *Reg Toxicol Pharmacol* 14:88-105.

Pope 2002. *Lung cancer, cardiopulmonary mortality, and long-term Exposure to fine particulate air pollution.* *JAMA* 287, 1132-1141

Liu, LJ, Koutrakis, P et al. (1993). *Use of Personal Measurements for Ozone Exposure Assessment: A Pilot Study.* *Environmental Health Perspectives* 101 (4) .

Krewski et al. (2000). *Reanalysis of the Harvard Six Cities Study and the ACS Study of Particulate Air Pollution and Mortality.* Special Report. HEI. Part 2,.

Krewski et al. (2003). *Overview of the Reanalysis of the Harvard Six Cities Study and American Cancer Society Study of Particulate Air Pollution and Mortality.* *J. Tox Env Hlth Part A*, 66: 1507-1551

Lipfert FW, Perry HMJ., et al. (2000) *The Washington University-EPRI Veterans' cohort mortality study: Preliminary results.* *Inhalation Toxicology* 12(Supp 4):41-73.

Lumley T and Sheppard L. (2000) *Assessing seasonal confounding and model selection bias in air pollution epidemiology using positive and negative control analyses.* *Environmetrics* 11:705-717.

Moolgavkar, (1995). *Air pollution and daily mortality in Philadelphia.* *Epidemiology* 6: 476-484

Moolgavkar SH. (2003) *Air pollution and daily mortality in two US counties:season-specific analyses and exposure-response relationships.* *Inhalation Toxicology* 15:877-907.

Moolgavkar SH and Luebeck EG. (1996) *A critical review of the evidence on particulate air pollution and mortality.* *Epidemiology* 7:420-428.

Nicolich et al. 1999. Evidence of a threshold effect for TSP in the Philadelphia data set. *J. Env. Med.* 1: 279-290.

OAQPS. (2003) Review of the national ambient air quality standards for particulate matter: policy assessment of scientific and technical information, OAQPS staff paper-first draft. Office of Air Quality Planning and Standards.

Pope CAI, Schwartz J, Ransom MR (1992) Daily mortality and PM10 pollution in Utah valley. *Arch Env Health* 47:211-217.

Pope CAI, Burnett RT., et al. (2002) Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *JAMA* 287:1132-1141.

Rabl 2003. Interpretation of air pollution mortality number of deaths or years of life lost? *JAWMA* 53: 41-50.

Roberts S. (2004) Biologically plausible particulate air pollution mortality concentration-response functions. *Environ Health Perspect* 112:309-313.

Samet JM, Zeger SL, Berhane K. (1995) Particulate Air Pollution and Daily Mortality: Replication and Validation of Selected Studies. Phase I Report, Health Effects Institute, Boston, MA.

Schlesinger, RB and Cassee, F. (2003). Atmospheric Secondary Inorganic Particulate Matter: The Toxicological Perspective As A Basis For Health Effects Risk Assessment. *Inhalation Toxicology* 15: 197-235.

Schwartz J. (1994) Air pollution and daily mortality: a review and meta-analysis. *Environ Res*, 64:36-52.

Schwartz J. (2004) Is the association of airborne particles with daily deaths confounded by gaseous air pollutants? An approach to control by matching. *Environ Health Perspect* 112:557-561.

Schins, R.P.F. (2004). Inflammatory effects of coarse and fine particulate matter in relation to chemical and biological constituents. *TAP* 195, 1-11.

Schwartz J. and Zanobetti A. (2000) Using meta-smoothing to estimate dose-response trends across multiple studies, with application to air pollution and daily deaths. *Epidemiology* 11:666-672.

Smith et al. 2000. Threshold dependence of mortality effects for fine and coarse particles in Phoenix, Arizona. *JAWMA* 50: 1367-1379.

Smith RJ, Davis JM, Speckman P. (1997) Assessing the human health risk of atmospheric particles. *ASA Proceeding Environmental Statisticians*.

Smith MT and Rothman N. (2000) Biomarkers in the Molecular Epidemiology of Benzene-exposed Workers. *J Toxicol Environ Health Part A* (61):439-445.

Sunyer 1996. Air pollution and mortality in Barcelona. *J. Epidem. Comm. Health* 50 (supple 1); S76-80.

Le Tertre, Median, S., Samoli, E, Forsberg, B, Michelozzi, P, Boumghar, A, Von, J, Bellini, A, Atkinson, R, Ayres, J, Sunyer, J, Schwartz, J, and Katsouyanni, K. (2002). Short-term effects of particulate air pollution on cardiovascular diseases in eight European Cities. *J. Epidemiology Community Health* 56:773-779.

Walton 2003. Is there a threshold for associations between ozone concentrations and health outcomes? A review. *AIRNET Rome*.

WHO (2004a). *Health Aspects of Air Pollution - Answers to Follow-Up Questions from CAFE*.

WHO (2004b) Anderson et al. *Meta-analysis of time-series studies and panel studies of particulate matter and ozone.*

**Reference Sub List 1 - A few examples of human clinical studies of ozone in potentially susceptible groups**

Drechsler-Parks, D.M. 1995. Cardiac output of O<sub>3</sub> and NO<sub>2</sub> exposure in healthy older adults. *Tox Ind. Hlth.* 11: 99-109

Gong, H. Hr., McManus, M.S., Linn, W.S. 1997. Attenuated response to repeated daily ozone exposures in asthmatic subjects. *Arch. Env. Hlth.* 52: 32-41.

Gong, H., Wong, R. Sarma, R.J., Sullivan, W.S., Shamoo, E.D., Anderson, K.R. and Prasad, S.B. 1998. Cardiovascular effects of ozone exposure in human volunteers. *Amer. J. Crit. Care Med.* 158: 538-546.

Horstman, D.H., Ball, B.A., Brown, J., Gerrity, T., Folinsbee, L.J. 1995. Comparison of pulmonary responses of asthmatics and non-asthmatic subjects performing light exercise while exposed to a low level of ozone. *Tox. Ind. Hlth.* 11: 369-385.

Kehrl, H. R., Hazucha, J. J., Solic, J. J., Bromberg, P. A. 1985. Responses of subjects with chronic obstructive pulmonary disease after exposures to 0.3 ppm ozone. *Amer. Rev. Resp. Dis.* 131: 719-724.

Koenig, J. Q., Covert, D.S., Smith, M.S., VanBelle, G., Pierson, W.E. 1988. The pulmonary effects of ozone and nitrogen dioxide alone and combined in healthy and asthmatic adolescent subjects. *Tox Env. Hlth* 4: 521-532.

Linn, W.S., Shamoo, D. A., Venet, T.G., Valencia, L.M., Anzar, U.T., Hackney, J.D. 1983. Response to ozone in volunteers with chronic obstructive pulmonary disease. *Arch. Env. Hlth.* 38(5): 278-283,

Linn, W.S., Fisher, D.A., Medway, D. A., Anzar, U.T., Spier, C. E., Valencia, L. M., Venet, T. G., and Hackney, J. D. 1982. Short-term respiratory effects of 0.12 ppm ozone exposure in volunteers with chronic obstructive pulmonary disease. *Am. Rev Resp. Dis* 125: 658-683.

Linn, W.S., Shamoo, D.A., Anderson, K.R., Peng, R.C., Avol, E.L., Hackney, J.D. 1994. Effects of prolonged, repeated exposure to ozone, sulfuric acid, and their combination in health and asthmatic volunteers. *Amer. J. Resp. Crit. Care Med.* 150: 431-440.

McBride, D.E., Koenig, J.Q., Luchtel, D.L., Williams, P.V., and Henderson, W. R. 1994. Inflammatory effects of ozone in the upper airways of subjects with asthma. *Am. J. Resp. Crit Care Med.* 149: 1192-7.

McDonnell, W.F. 1989. Individual variability in the magnitude of acute respiratory responses to ozone exposure. In M.J. Utell and R. Frank, editors. *Susceptibility to Inhaled Pollutants.* ASTM, Philadelphia. 75-78.

Superko, H.R., Adams, W.C., Daly, P.W. (1984). Effects of ozone inhalation during exercise in selected patients with heart disease. *Amer. J. Med.* 77: 462-470.

U.S. Environmental Protection Agency. 1988. *Summary of selected new information on effects of ozone on health and vegetation: Draft supplement to Air Quality Criteria for Ozone and other Photochemical oxidants.* EPA/8-88/105A, ECAO, Research Triangle Park, North Carolina.

**Attachment 3**

**PROPOSAL FOR IMPROVED LAYOUT OF CAFE CBA COMMUNICATION**

**Concern:**

Human Health

<b>Health end point in 2002</b>	<b>Expected improvement in 2020<sup>1</sup></b>
<b>Particulate Matter</b>	
3 million life years lost annually <sup>2</sup>	Reduction by 1 million <sup>2</sup>
80,000 hospital admissions annually	Reduction by 30,000
<b>Ground Level Ozone</b>	
x thousand life years lost annually <sup>3</sup>	Reduction by y

<sup>1</sup>The reader should be aware that it will not be possible to track actual improvements based on individual level data, as is the case for safety statistics. All estimates for air pollution are based on past studies that relied upon group measures of exposure and effect.

<sup>2</sup>Estimates are based on non-EU studies and assume that even close to zero exposure to man-made particulate matter will result in deaths.

<sup>3</sup>Estimates assume that both high and low levels of ozone, down to background levels, produce mortality.

In addition to these possible improvements, there may be improvements in other health endpoints but these estimates are even more uncertain than those mentioned above. These endpoints include a possible reduction in respiratory medication use days with reduced levels of particulate matter and ozone, and a reduced number of restricted activity days with reduction in particulate matter.

**Response:**

We have commented on the general principles, above. While there are some good presentation ideas in the table, we also find serious limitations for several reasons, the most important being:

- The scope of this table is very limited.
- Uncertainty can lead to underestimation – this Table implies that overestimation is the only issue to be considered.